

BERCEO	118-119	83-93	Logroño	1990
--------	---------	-------	---------	------

## APLICACIÓN DE DOS TÉCNICAS MULTIVARIANTES (CLUSTER NO JERARQUIZADO Y ANÁLISIS DE COMPONENTES PRINCIPALES) A UN ESTUDIO DE GEOGRAFÍA ELECTORAL: EL VOTO EN LA CIUDAD DE LOGROÑO Y SUS CONDICIONANTES SOCIALES\*

Florencio Ortiz Alejos\*\*  
Ignacio Sobrón García\*\*

### RESUMEN

*Con este trabajo se pretenden resumir algunas de las posibilidades que ciertas técnicas de análisis estadístico multivariante abren al estudio de la Geografía Urbana. Concretamente, se aborda el problema de la distribución del voto en la ciudad de Logroño y se establece la relación que presenta con una serie de variables socioeconómicas y demográficas que lo condicionan sensiblemente. Los autores han empleado básicamente para establecer tales relaciones dos técnicas de análisis multivariante: sistemas de "clustering" no jerarquizadas y el análisis factorial vía componentes principales. Cada una de ellas mostrará una serie de rasgos específicos de interés para el establecimiento de la cuestión considerada: por una parte, la realización de una cartografía temática adecuada al problema y la definición de áreas sociopolíticas homogéneas y, por otra, la formulación de una serie de relaciones nítidas entre voto y variables socioeconómicas y demográficas.*

*Palabras clave: geografía electoral, técnicas multivariantes, Logroño.*

*In this paper some of the possibilities of certain procedures of multivariate statistics on vote prediction and Urban Geography are discussed. Attention is paid mainly to the problem of vote distribution in the city of Logroño (La Rioja, Spain) and relationships between some social variables and vote are described. The authors have achieved all this using two current techniques of multivariate analysis: non hierarchical clustering and Principal Component Analysis. Each of them will result in a series of specific traits which would help in the knowledge of the mentioned question: first, monitoring and thematic cartography of the data and definition of homogeneous socio-politic areas and, second, the formulation of a series of clear relationships between vote and some socio-demographic variables.*

*Key words: Electoral Geography, multivariate statistics, Logroño.*

\* Comunicación presentada al II Coloquio sobre Geografía de La Rioja.  
Recibido en Junio de 1989. Aprobado el 14 de Septiembre de 1989.

\*\* Investigador Agregado. Instituto de Estudios Riojanos.

El análisis de los resultados electorales de la ciudad de Logroño posee un interés indudable para el conocimiento de la realidad sociopolítica regional, consecuencia de la concentración en la capital de algo más del 42% del conjunto electoral regional (ORTIZ, 1988). Por otra parte, la existencia de datos sociodemográficos a nivel de sección censal posibilita el establecimiento de una serie de relaciones específicas entre el voto y sus condicionantes socioeconómicos y demográficos.

El presente trabajo aborda de modo concreto el análisis de la distribución del voto en la ciudad de Logroño y establece algunos de los principales factores que lo condicionan sensiblemente. Se trata, a partir de una aproximación de carácter geográfico-ecológico urbano, de explicar en función de las estructuras sociales y las características demográficas de la población, los diversos comportamientos del electorado logroñés. Se prestará especial atención a las cuestiones relativas a la representación cartográfica de los datos, tratando de mostrar de un modo gráfico las principales relaciones entre estructura sociodemográfica del electorado y su comportamiento electoral.

El principal problema que presenta cualquier trabajo de este tipo es la relativa escasez y dispersión de las fuentes estadísticas existentes. En este caso concreto, se han podido considerar algunas variables potencialmente significativas desde el punto de vista electoral, como edad, sexo, procedencia, nivel de instrucción, situación laboral y nivel profesional; sin embargo, no han podido ser considerados aspectos de tanto interés como nivel de práctica religiosa, características de la vivienda, nivel de renta, etc. debido a que la información existente no se halla codificada a nivel de sección censal.

## 1.- FUENTES Y UNIDADES DE ANÁLISIS

Los datos electorales empleados han presentado una dispersión notable, dada la carencia de un organismo que los centralice, creando a veces problemas de falta de coincidencia de menor interés. Las variables electorales son las referidas al voto para el Congreso de los Diputados en las elecciones legislativas de 1982, escogidas por su representatividad, ya que a partir de ellas se produce un cambio esencial en la estructura electoral logroñesa, coincidente con uno global a nivel nacional y tendente al bipartidismo. Otro factor que nos ha condicionado en la elección de estos comicios ha sido el bajo nivel de abstención (15% aprox.), el más bajo de todas las elecciones celebradas en Logroño, tal vez como consecuencia del interés despertado por los cambios políticos y la polarización electoral en torno a dos partidos<sup>1</sup>.

Las variables sociodemográficas utilizadas han sido extraídas del Padrón Municipal de Habitantes de Marzo de 1981 mediante un muestreo aleatorio y sistemático de un 20% del total de la población de cada sección censal. El padrón de 1986 no ha podido ser utilizado por la carencia de ciertos datos de inexcusable interés como profesión y nivel profesional de los activos.

La unidad de análisis considerada ha sido la sección censal, al ser la unidad espacial de menor tamaño para las cuales existen datos. El problema más grave que presenta este método es la falta de homogeneidad socioeconómica o demográfica en su interior, que aún no

<sup>1</sup> Datos referentes a otras consultas electorales, desde 1977 hasta 1987, pueden ser consultado en ORTIZ ALEJOS (1988), donde se describe con precisión la distribución del voto y sus condicionantes en todas las elecciones del actual período democrático.

siendo demasiado importante, puede enmascarar o desfigurar en ciertos casos algunos comportamientos electorales determinados.

## 2.- MÉTODOS DE TRABAJO

Una vez obtenidos los datos sociodemográficos y electorales a nivel de sección censal se ha procedido básicamente a su organización y tratamiento estadístico. Como paso previo, se ha realizado un análisis de correlaciones entre los datos de los diversos partidos entre sí y de éstos con las diferentes variables sociodemográficas. Con esta labor se trata de comprobar básicamente las relaciones más simples existentes entre el voto y sus condicionantes socioeconómicos y de tratar de prever cuáles pueden ser los principales factores que explican el comportamiento electoral en la ciudad de Logroño.

Aunque no serán presentados en este trabajo, por problemas de extensión, se han establecido modelos de predicción lineal del voto a partir de predictores socioeconómicos, basados esencialmente en ecuaciones de regresión múltiple paso a paso y en diferentes tests de significación adecuados al problema (SEARLE, 1971; WINER, 1971; NETER *et al.*, 1974; COHEN *et al.*, 1983; BENDEL *et al.*, 1977). Empleando como variables independientes 23 características sociodemográficas se han planteado ecuaciones de regresión que eliminan los predictores menos interesantes y los que presentan altos grados de correlación, de modo que se evita el problema de la colinealidad. De este modo, empleando modelos que incluyen aproximadamente 7 predictores sociodemográficos, se ha logrado predecir el voto de los principales partidos políticos con una aproximación que oscila entre el 89% y 92% para PSOE y AP y en torno al 75-80% para los datos del PCE o de la abstención. Creemos que estos resultados pueden ser mejorados, y se tiene previsto hacerlo, mediante la consideración de un análisis de series temporales que permita incorporar fenómenos evolutivos al modelo predictivo sociodemográfico.

Uno de los aspectos a los que más interés hemos dedicado ha sido al empleo del análisis factorial de componentes principales, con el fin de establecer su potencialidad para un estudio de geografía electoral. Esta técnica ya había sido empleada por uno de los autores y aplicada al establecimiento de vectores explicativos de la organización social de la ciudad de Logroño (ORTIZ, 1989). En el caso que nos ocupa, el P.C.A., aparte de como síntesis multivariante de la información aportada por métodos de correlación más simples, nos va a permitir realizar una primera aproximación cartográfica adecuada a la problemática de las relaciones voto-factores sociodemográficos y un establecimiento de las principales relaciones entre el voto y los factores más ligados al problema de su distribución. Resulta de un gran interés la representación cartográfica de las secciones censales de acuerdo a la distribución de cada una de las observaciones en relación a cada uno de los nuevos vectores generados. En lugar de emplear la típica representación de las observaciones en un eje de coordenadas (bi- o tri-dimensional), hemos preferido su cartografía espacial sobre una trama de la ciudad de Logroño elaborada a nivel de sección censal, puesto que de este modo se reflejan las principales tendencias espaciales de la relación conjunta voto-factores que lo condicionan.

Este procedimiento parece bastante adecuado al problema planteado, debido a la estructura típica del análisis de componentes (HARMAN, 1980; CUADRAS, 1981), al determinar una serie de combinaciones lineales empíricas de las variables consideradas que expliquen la mayor parte de la varianza observada, ante todo en conjuntos de datos, como el que ahora nos ocupa, con medidas de gran variación y medidas en las mismas unidades; de

este modo se logran evitar los problemas de distorsión de los elipsoides cuyo eje se corresponde con los factores principales.

Se han realizado cinco análisis de este tipo con diferente número de variables en cada uno con el fin de obtener los mejores resultados. En el modelo que presentamos se han incluido 30 variables referentes a estructura por edad y sexo de la población, procedencia y antigüedad de los inmigrantes, nivel de instrucción, situación laboral, nivel profesional y datos ligados al voto a los principales partidos y a la abstención en la convocatoria de 1982. El número de unidades de observación computadas ha sido de 81, correspondientes al número de secciones censales establecidas por el Ayuntamiento en dicho momento.

Se han seleccionado tan sólo aquellos componentes cuyo autovalor es superior a 1.5 y que explican más de un 5% de la varianza. Asimismo, se ha realizado una rotación de ejes de tipo Varimax para mejorar la interpretación de los ejes. Posteriormente se han seleccionado aquellos componentes con un claro valor sociopolítico, desechándose los demás, y se han analizado los pesos (Scores) de cada factor en cada unidad de observación, a través de los cuales se va a proceder a una cartografía que muestre la distribución espacial de los componentes con significación electoral. En nuestro caso, se ha escogido el eje 1 y sobre una trama a nivel de sección censal de la ciudad de Logroño, se ha representado la posición de cada observación con respecto al eje mencionado. Se ha preferido este tipo de representación a la tradicional de los ejes en disposición cartesiana por las posibilidades que ofrecía el método para una cartografía directa del espacio considerado, pudiéndose establecer nítidas relaciones espaciales entre zonas y observaciones.

Con el fin de llegar a la definición de áreas sociopolíticas homogéneas y de proceder a una cartografía-resumen, lo más significativa posible, de los datos, los autores han empleado otra técnica de análisis estadístico multivariante que se ha mostrado enormemente adecuada para el tratamiento de la cuestión considerada. Esta ha sido la técnica de agrupamiento no jerarquizado o "clustering no jerarquizado". Este método se diferencia sensiblemente del cluster jerárquico o de dendrograma y puede ser considerado como un método de gran interés para el estudio y clasificación de grupos de datos distribuidos de modo disperso en el espacio, logrando establecer lazos de unión entre las diversas observaciones, de tal modo que se puede proceder a una clasificación de los datos sobre el mapa, basada en los principales factores diferenciales y que muestre al mismo tiempo las principales pautas de variación y relación espacial a nivel de grupo. Este proceso (descrito en algunas obras, entre las que destaca HARTIGAN, 1975) se basa ante todo en la separación de los miembros de un grupo estadístico en un número determinado de grupos, considerando las variaciones tanto dentro de los grupos como entre ellos. Las diferentes observaciones se van reasignando entre los diferentes grupos mediante un procedimiento semejante en ciertos aspectos a un método clasificatorio de máxima probabilidad, llegándose al resultado final tras una serie de iteraciones suficientes como para asignar correctamente las observaciones a los grupos. Con este procedimiento llegamos a establecer cuáles son los principales factores que generan la separación entre los datos y cuáles deben ser desechados y, a partir de ellos, proceder a la creación de grupos y a la asignación de las observaciones a cada uno de ellos.

Los datos empleados son los referentes, en primer lugar, a aspectos sociodemográficos, con el fin de proceder a una diferenciación de áreas sociales homogéneas en Logroño; a continuación, se ha realizado un cluster con variables electorales con el fin de comparar las semejanzas y diferencias entre la distribución social y la del voto. Los resultados obtenidos aparecen reseñados en los cuadros 2 y 3 y han sido posteriormente cartografiados en

los gráficos 2 y 3. Se han realizado varias pruebas con diferente número de grupos y de iteraciones, hasta llegar a los resultados presentados, que reducen el número de clusters a 6, número suficiente para mostrar diferencias significativas a nivel de grupo, pero que, sin embargo, evita los problemas de distorsión y ruido generados al emplear mayor número de agrupaciones cluster.

### 3.- RESULTADOS Y DISCUSION

Los principales resultados obtenidos aparecen resumidos en los cuadros y gráficos adjuntos. En primer lugar, puede referirse, que, mediante el empleo del PCA, se consigue mostrar la presencia de un fuerte componente socioelectoral en la ciudad de Logroño, que reúne por una parte la asociación voto de derecha-nivel socioeconómico-cultural alto, frente a los factores ligados al voto de los partidos de izquierda, que coinciden con niveles socioculturales más bajos. Al cartografiar espacialmente los datos de este análisis (gráfico 1) se establece una zonificación de la ciudad de Logroño en base a este eje sociopolítico, mostrándose una estructura de tendencia centrífuga y semiconcéntrica alterada tan sólo por algunas polaridades direccionales. Se establece una clara gradación sociopolítica desde el centro hacia la periferia.

Esta diferenciación espacial a nivel social y político queda claramente reflejada en los resultados de los análisis cluster, puesto que definen nítidamente áreas de semejante comportamiento sociodemográfico y electoral, caracterizadas por diferentes valores de las variables consideradas. Nuevamente se aprecia una distribución de ambos fenómenos (socio-demográfico y electoral) semejante a la ofrecida por el PCA, aunque ahora aparece más claramente reflejada la compacidad de los grupos determinados y alguna de sus tendencias espaciales, por ejemplo, la prolongación hacia el sur de las zonas de características socioeconómicas más elevadas y su consiguiente reflejo electoral. Por otra parte, al cartografiar los datos (gráficos 2 y 3) podemos llegar a la realización de una serie de mapas temáticos de indudable interés para conocer la realidad sociopolítica logroñesa, estableciendo una zonificación de la ciudad de acuerdo a los factores considerados. Las relaciones entre voto y estructura sociodemográfica, claramente señaladas a partir del PCA, pueden ser observadas a nivel de grupos homogéneos en los resultados de los análisis cluster (cuadros 2 y 3) y resulta de notable interés, por su enorme grafismo, la comparación de los dos mapas que muestran áreas sociales y electorales homogéneas, apreciándose una enorme correspondencia (con variaciones tan sólo individuales y de carácter marginal) entre ambos y reflejando claramente la presencia de unos condicionantes sociodemográficos en la distribución del voto logroñés.

### REFERENCIAS BIBLIOGRÁFICAS

- BENDEL, R.B. y AFIFI, A.A., 1977.- "Comparision of stopping rules in forward stepwise regression". *Jour. Amer. Stats. Ass.*, 72:46-53.
- COHEN, J. y COHEN, P., 1983.- *Applied multiple regresion/corelation analysis for the behavioural sciences*. Hillsdale Ed., N.J. Lawrence Erlbaum.
- CUADRAS, C.M., 1981.- *Métodos de análisis multivariante*. Universidad Barcelona, 642 pp.
- HARMAN, H.H., 1980.- *Análisis factorial moderno*. Saltés, Madrid, 531 pp.
- HARTIGAN, J.A., 1975.- *Clustering Algorithms*. John Wiley and Sons, New York.

- NETER, J., WASERMAN, W. y KUTNER, M.H., 1985.- *Applied linear statistic models*. Richard D. Irwing Inc., Homewood, 111 pp.
- ORTIZ ALEJOS, F., 1988.- *Geografía electoral de Logroño (1977-1987)*. Instituto de Estudios Riojanos, Logroño, 361 pp.
- ORTIZ ALEJOS, F., 1989.- *La diferenciación social del espacio urbano de Logroño*. Instituto de Estudios Riojanos - Ayuntamiento de Logroño, Logroño, 117 pp.
- SEARLE; S. R., 1971.- *Linear Models*. John Wiley and sons, New York.
- WINER, B.J., 1971.- *Statistical principles in experimental design*. McGraw & Hill, New York.

**Cuadro 1**  
**Varianza explicada por los distintos componentes**

Componente	Eigenvalue	Varianza absorbida	Varianza acumulada
1	7,81	26,05	26,05
2	4,59	15,31	41,36
3	3,70	12,34	53,70
4	2,68	8,95	62,65
5	2,77	9,26	71,91

**Las principales saturaciones en los cinco componentes**

COMPONENTE 1.	MUJERES	-0,48	SININSTR.	0,45
	NIMEDIO	-0,86	NIELEMEN	0,82
	NISUPERI	-0,89	NPMEDBAJ	0,77
	NPALTO	-0,85	NPBAJO	0,72
	NPMEDIO	-0,74	PARO	0,48
	AP	-0,83	PSOE	0,82
	UCD	-0,38	PCE	0,46
COMPONENTE 2.	ADMADURO	-0,81	ADJOVEN	0,84
	ANCIANOS	-0,84	INMI 71-80	0,76
	JUBILADOS	-0,87		
	PARO	-0,46		
COMPONENTE 3.	INEXPRO	-0,89	INMI 61-70	0,67
	SININSTR	-0,46	INMPRO	0,87
	ABS	-0,57		
	FN	-0,48		
COMPONENTE 4.	AMASCASA	-0,76	INMIGRAC	0,72
	PCE	-0,45	MUJERES	0,42
			PAFEMENI	0,59
COMPONENTE 5.	PACTIVA	-0,92	SININSTR	0,47
	PAFEMENI	-0,47		
	PACTMASC	-0,81		

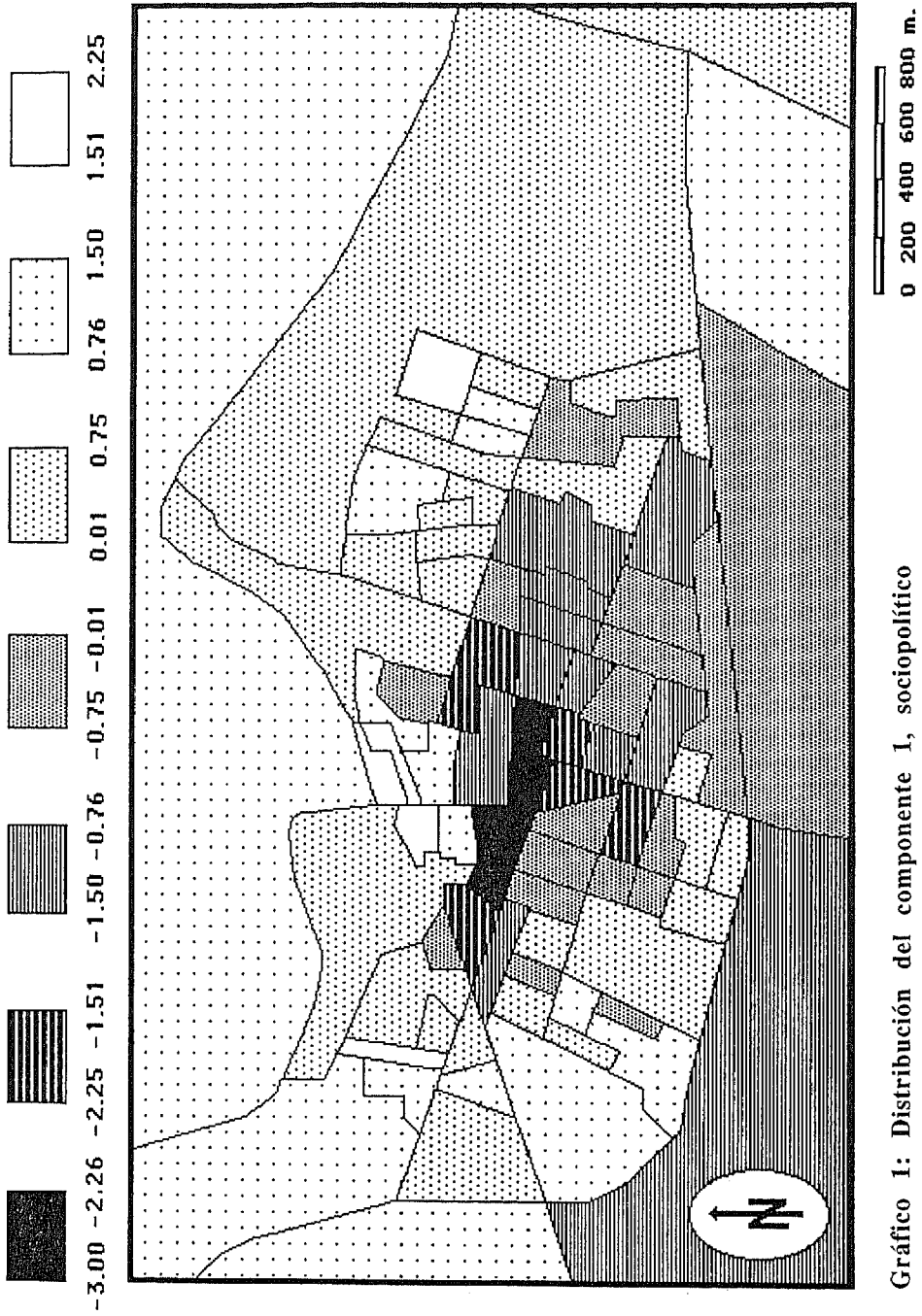


Gráfico 1: Distribución del componente 1, sociopolítico

**Cuadro 2: Resumen del análisis cluster realizado con variables sociodemográficas**

Variables	Cluster nº 1		Cluster nº 2		Cluster nº 3		Cluster nº 4		Cluster nº 5		Cluster nº 6	
	Media	$\sigma$	Media	$\sigma$	Media	$\sigma$	Media	$\sigma$	Media	$\sigma$	Media	$\sigma$
Adultos jóvenes	46.14	5.11	41.76	3.49	32.94	0.00	33.12	0.00	36.67	4.53	40.65	3.09
Adultos maduros	19.33	5.33	22.38	3.26	31.35	0.00	27.92	0.00	24.53	3.20	23.82	3.00
Ancianos	7.49	2.64	11.37	3.95	19.05	0.00	20.78	0.00	14.93	4.92	14.66	3.45
Mujeres	50.24	1.88	52.57	1.62	45.63	0.00	50.85	0.00	51.49	2.01	53.45	1.83
Sin instrucción	2.85	2.77	1.70	1.50	9.52	0.00	12.69	0.00	7.71	5.30	1.66	1.01
Instrucción elemental	82.80	4.94	74.48	3.96	73.81	0.00	79.85	0.00	79.82	5.03	61.14	5.36
Instrucción media	9.46	2.72	12.79	2.26	16.67	0.00	3.73	0.00	7.23	2.41	17.32	3.07
Instrucción superior	4.54	2.37	10.81	2.83	0.00	0.00	3.73	0.00	4.22	1.96	19.40	3.80
Población activa	37.25	3.16	36.75	1.72	32.54	0.00	34.56	0.00	34.83	2.68	35.38	1.59
Población activa femenina	19.92	3.03	21.69	1.81	11.30	0.00	17.57	0.00	18.93	2.80	22.15	1.20
Paro	11.75	2.54	10.94	2.53	17.07	0.00	24.71	0.00	16.05	5.42	10.98	1.44
Nivel profesional alto	8.00	4.53	19.11	4.63	0.00	0.00	9.50	0.00	8.01	4.65	35.03	4.69
Nivel profesional medio	23.71	4.91	35.38	4.93	60.00	0.00	19.00	0.00	23.63	7.00	36.12	4.10
Nivel profesional medio-bajo	51.17	6.55	33.55	6.13	33.30	0.00	23.70	0.00	44.37	5.56	20.34	5.22
Nivel profesional bajo	17.19	5.36	12.00	4.44	6.70	0.00	47.60	0.00	23.99	6.87	8.75	3.74



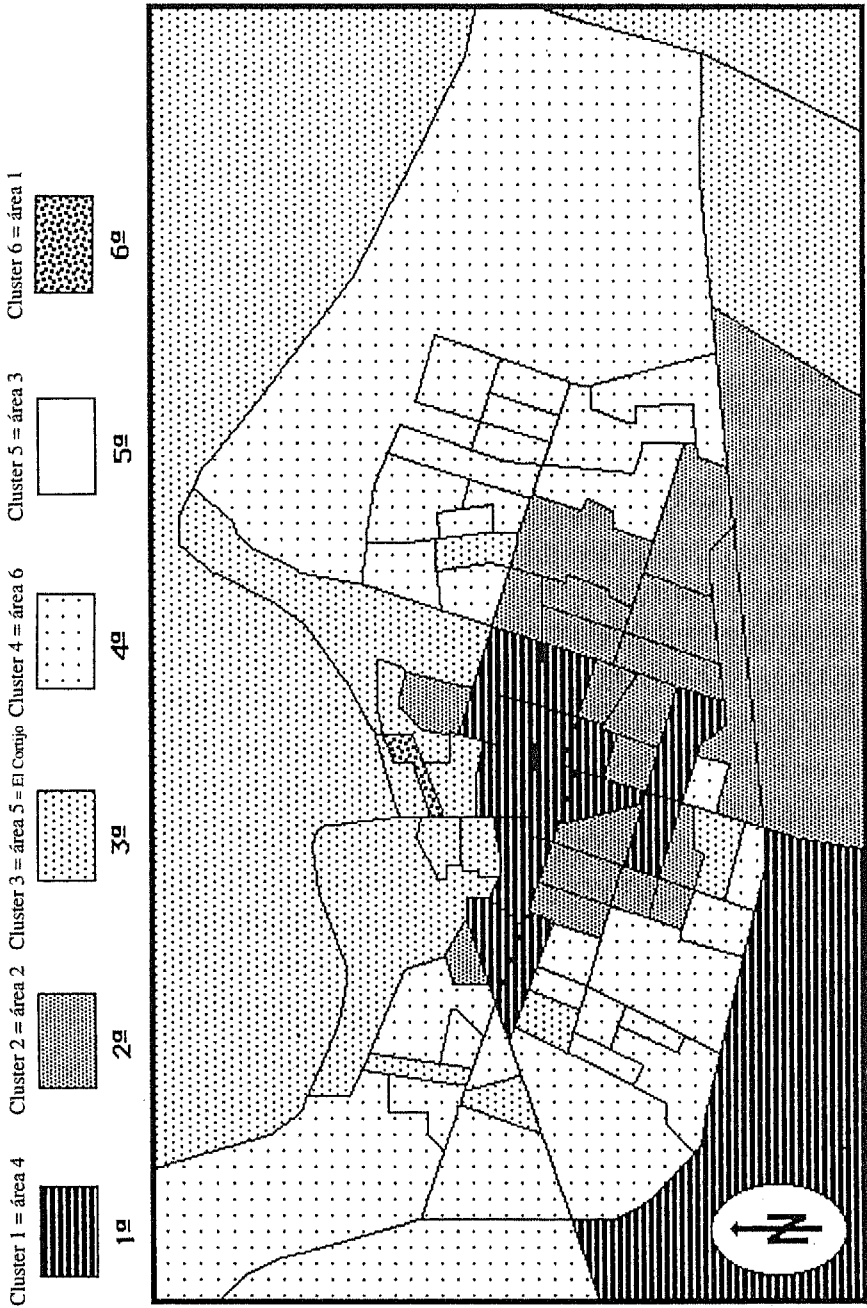


Gráfico 2: Logroño. Áreas sociales homogéneas  
Correspondencia entre clusters y áreas sociales homogéneas

Cuadro 3: Resumen del análisis cluster realizado con los resultados de las elecciones legislativas de 1982 (Congreso)

Variables	Cluster nº 1		Cluster nº 2		Cluster nº 3		Cluster nº 4		Cluster nº 5		Cluster nº 6	
	Media	$\sigma$	Media	$\sigma$	Media	$\sigma$	Media	$\sigma$	Media	$\sigma$	Media	$\sigma$
P.S.O.E.	47.27	3.03	39.33	2.72	45.45	0.00	27.47	4.67	44.57	1.28	57.14	3.52
A.P.	37.15	2.52	45.30	2.55	27.27	0.00	57.75	4.92	40.39	2.04	26.60	3.61
F.N.	0.41	0.36	0.35	0.23	0.23	0.00	1.09	1.27	0.62	0.55	0.21	0.29
C.D.S.	3.90	0.82	4.14	0.89	2.73	0.00	3.58	1.07	3.73	1.11	3.58	0.90
U.C.D.	5.53	1.40	6.14	1.31	7.50	0.00	6.59	1.39	5.03	1.72	4.95	1.25
P.C.E.	1.50	0.61	1.19	0.55	5.45	0.00	0.83	0.45	1.86	0.24	2.31	1.34
Abstención	14.44	2.46	13.85	1.64	50.51	0.00	14.06	1.90	26.30	3.81	13.73	4.08

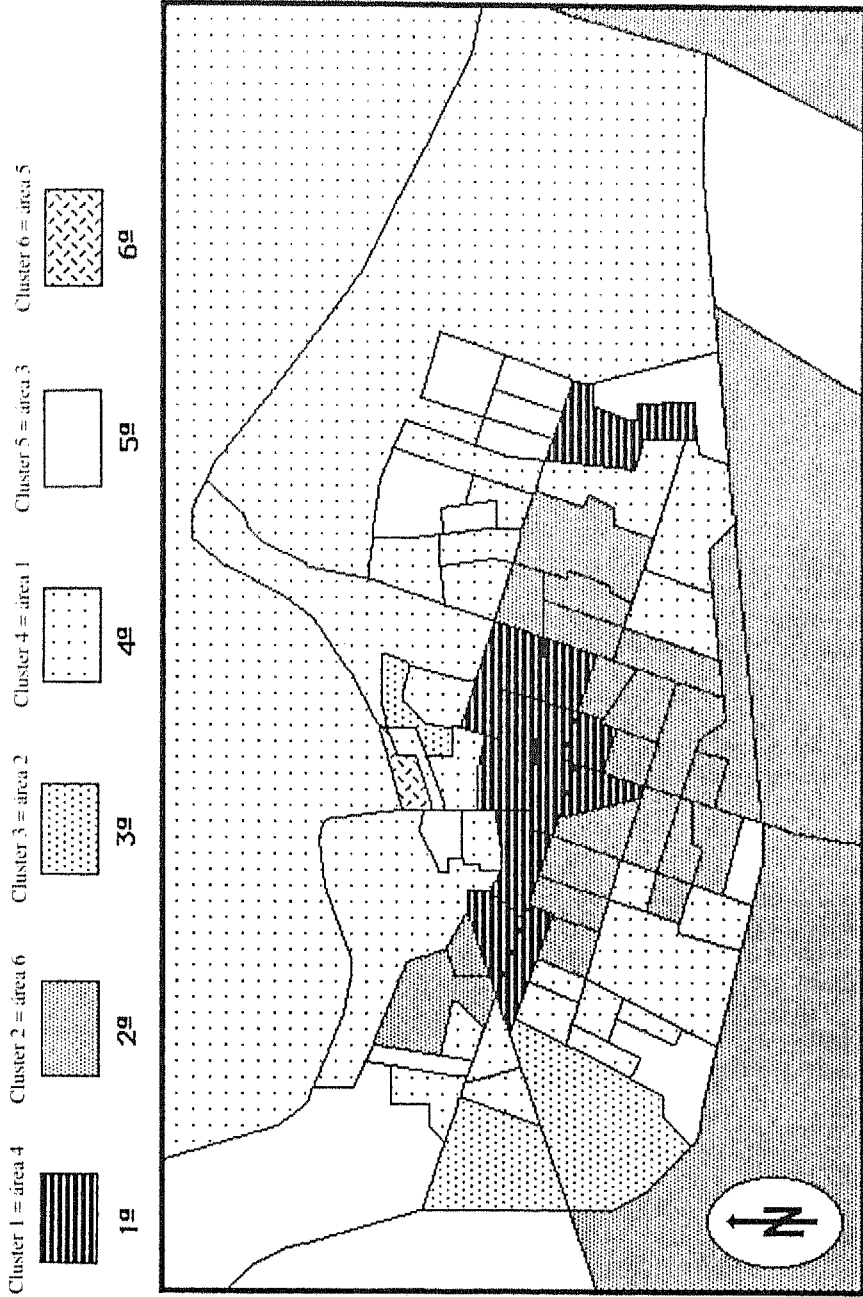


Gráfico 3: Logroño. Áreas homogéneas electorales en las elecciones legislativas de 1982  
Correspondencia entre clusters y áreas electorales homogéneas